63-3-5

MEMORANDUM
RM-3504-PR
MAY 1963

IMPROVING THE QUANTIZATION OF RANDOM SIGNALS BY DITHERING

G. G. Furman

PREPARED FOR:

UNITED STATES AIR FORCE PROJECT RAND

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-3504-PR
MAY 1963

IMPROVING THE QUANTIZATION OF RANDOM SIGNALS BY DITHERING

G. G. Furman

This research is sponsored by the United States Air Force under Project RAND—contract No. AF 49(638)-700 monitored by the Directorate of Development Planning, Deputy Chief of Staff, Research and Development, Hq USAF. Views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of the United States Air Force.

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

PREFACE

In processes involving analog machines, digital machines, or a combination of these, the accurate transmission of information is of great importance. This is sometimes aided by the application of independent quantizer activators called dithers. The present Memorandum gives a numerical solution to the practical problem of determining the effects of such activators upon the statistical processing properties of a quantizer.

The author is a RAND consultant. The work was initiated while at RAND during the summer of 1962 and continued under a Tri Service Contract Grant AFOSR-62-340 and a Raytheon Predoctoral Fellowship.

SUMMARY

This report gives a numerical solution to the practical problem of determining the effects of independent quantizer activators called dithers upon the statistical processing properties of the quantizer. For the highly important sinusoidal and sawtooth dithers exact analysis yields for the first time answers, as functions of dither amplitude, to the question of what upper bounds does the dither impose on the following: (1) correlation between the quantizer input and quantization noise, (2) value of the noise mean square and (3) fidelity in the transmission of the mean square, mean fourth and other even moments of the input. The above information, which also comprises a theorem for the quantization of sinusoids and sawtooths, indicates that the rarely used sawtooth is superior to the sinusoid.

CONTENTS

IMPROVING THE QUANTIZATION OF RANDOM SIGNALS BY DITHERING

1. INTRODUCTION

The injection of dither is a practical means of improving quantizer performance economically because it enables coarse quantizers to emulate ultrafine ones. Some indication of the many processes employing quantization was given in an earlier report [1], where the linearizing effect of dither upon the multistep quantizer nonlinearity Q (Fig. 1) was studied with the goal of minimizing the maximum excursion of the equivalent quantizer gain from perfect linearity. While such a minimization is useful for the design of systems which are subject to deterministic inputs, it does not solve the problem of optimizing the performance of the quantizer in systems which are subject to random inputs. For the random case, we shall examine the effects of dither upon the performance of the quantizer as a statistical operator, i.e., as an operator upon the moments of its input and the quantizer as a source of noise which can exhibit varying degrees of correlation with the input.

Consider the problem as formulated in Fig. 2 to be the following:

(i) Given that:

a. Transfer functions A and C may be linear or nonlinear and sampled data or time continuous, but they are

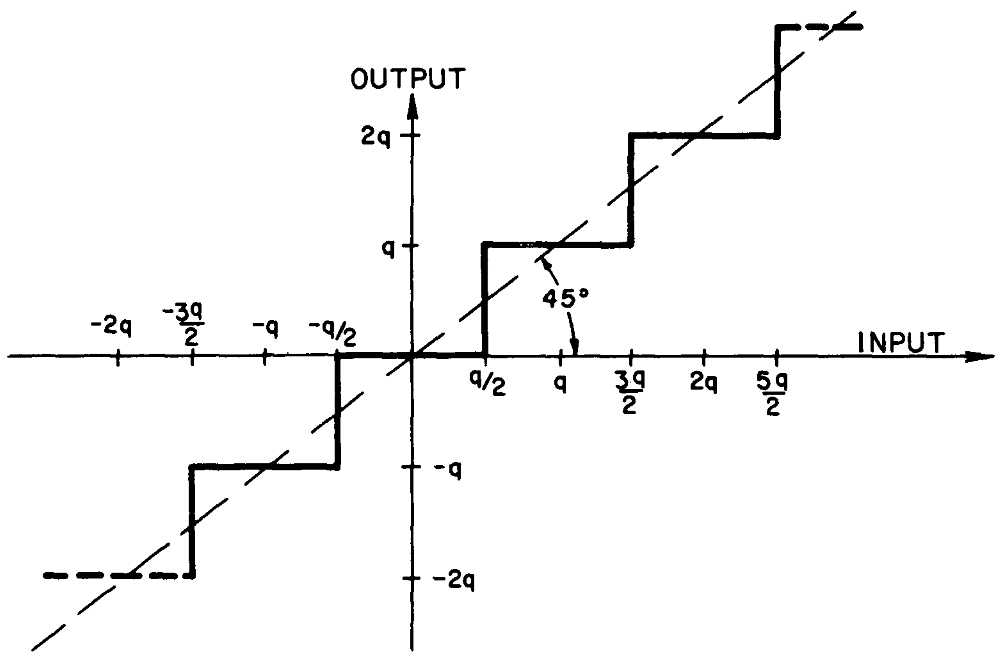

Fig. 1 -- The multistep quantizer nonlinearity, Q

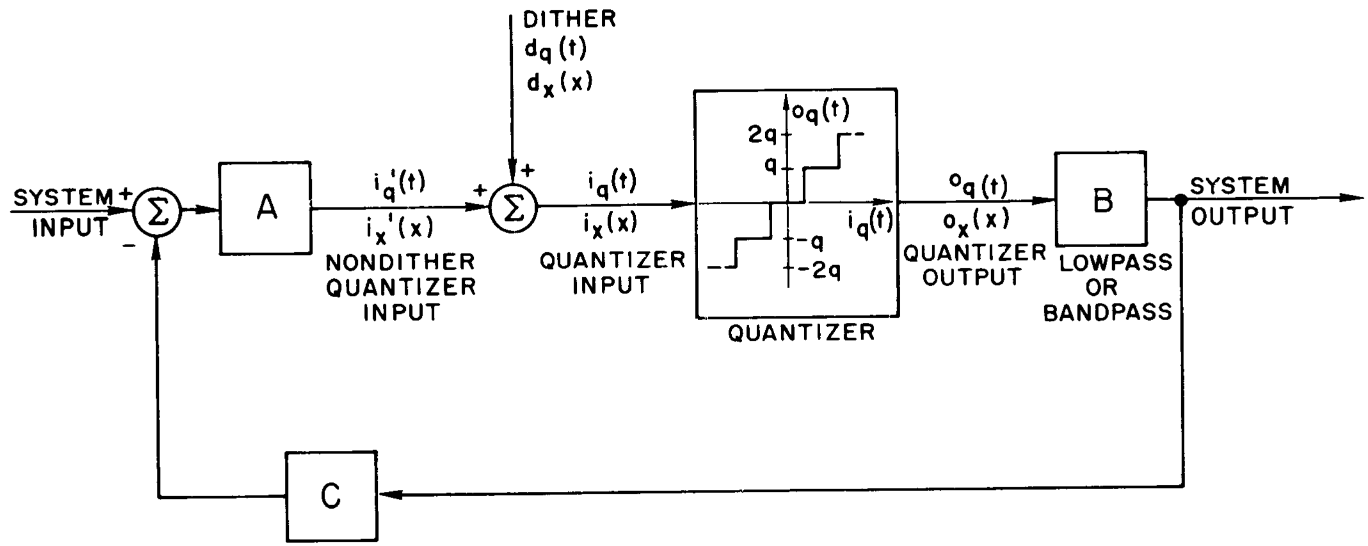

Fig. 2 -- Problem formulation

otherwise specified. It is possible to have $A = 1$, $C = 0$.

b. Transfer function B is linear and has either a low-pass or band-pass frequency characteristic. Also, if B is sampled-data, the sampling rate is high, the information content of $o_q(t)$ not being appreciably degraded by sampling.

c. A time periodic dither $d_q(t)$ may be injected in combination with the nondither quantizer input $i_q'(t)$. The quantizer output $o_q(t)$ consists of the total quantizer input $i_q(t)$ plus a quantization noise $n_q(t)$:

$$(1) \qquad o_q(t) = i_q(t) + n_q(t)$$

d. d_q may be sinusoidal (i.e., $d_q(t) = d(t)_{\sim} \equiv mq \sin \omega_d t$) or it may be sawtooth (i.e., $d_q(t) = d_{\Delta}(t) \equiv mq(1 - 2t/T_d)$ for $0 < t < T_d$), where q is the quantum size, m is the normalized dither amplitude and $\omega_d = 2\pi/T_d = 2\pi f_d$ is the dither radian frequency.

e. The variables i_q', d_q, i_q and o_q have the corresponding probability density functions (p.d.f.'s) $i_x'(x)$, $d_x(x)$, $i_x(x)$ and $o_x(x)$.

f. It is possible to operate with ω_d in the attenuation region of B, far above the frequency band containing the spectrum of i_q'. In this way much of the dither and noise does not reach the system output for n_q becomes nearly periodic, with ω_d as its fundamental.

Under certain input conditions [2, 3], the quantizer acts solely as a source of uniformly distributed noise which is uncorrelated with the quantizer input. The quantization is then said to be ideal. Our problem is, for the condition $i_q' = 0$, to develop measures of how ideal the quantization is. This furnishes quantization theorems for sinusoidal and sawtooth dithers. Also, if i_q' and d_q are uncorrelated, the measures for the condition $i_q' = 0$ can be used to develop bounds on the measures for the condition $i_q' \neq 0$, $d_q \neq 0$. This can be done with i_q' completely unknown, providing the conditions under (i) prevail.

(ii) The problem is to find as functions of m the following:

a. With $i_q' = 0$,

1. The zero shift correlation coefficient ρ_{dn} between d_q and n_q

$$\rho_{dn}(m) \equiv \frac{\overline{d_q\, n_q}}{\sigma_d\, \sigma_n} \tag{2}$$

where σ_d and σ_n are the standard deviations of d_q and n_q respectively.

2. The difference between $\overline{(n_q)^2}$ and $\overline{(n_{qu})^2} = q^2/(12)$, where n_{qu} has the p.d.f. $n_{ux}(x)$ such that

$$(3) \qquad n_{ux}(x) \equiv 1/q \quad \text{for} \quad |x| \leq q/2$$

$$\equiv 0 \quad \text{otherwise.}$$

The normalized excess noise EN is then

$$(4) \qquad EN(m) \equiv \frac{\overline{(n_q)^2}}{q^2} - \frac{1}{12} \ .$$

3. The difference between the actual and ideal values of $\overline{(o_q)^k}$ normalized with respect to the ideal kth moment. This dither figure of merit, DFM_k, is:

$$(5) \qquad DFM_k(m) \equiv \frac{\overline{(o_q)^k} - \overline{[d_q(m) + n_{qu}]^k}}{\overline{[d_q(m) + n_{qu}]^k}}$$

where $\overline{d_q n_{qu}} \equiv 0$.

b. With $i_q' \neq 0$, find as functions of m and i_q' the following:

$$(6) \qquad \rho_{in}'(i_q',m) \equiv \frac{\overline{i_q n_q}}{\sigma_i \sigma_n}$$

$$(7) \qquad EN'(i_q',m) \equiv \frac{\overline{(n_q)^2}}{q^2} - \frac{1}{12}$$

$$(8) \qquad DFM_k'(i_q',m) \equiv \frac{\overline{(o_q)^k} - \overline{[i_q + n_{qu}]^k}}{\overline{[i_q + n_{qu}]^k}}$$

where σ_i is the standard deviation of i_q and $\overline{i_q n_{qu}} \equiv 0$.

Solutions for Eqs. (6), (7), and (8) are possible if $i_q'(t)$ is known. However, without this knowledge it is still possible to solve for the least upper bounds of $|\rho_{in}'|$, $|EN'|$ and $|DFM_k'|$ by taking the worst possible case, i.e., $i_q'(t) = c$, a constant. For any measure, say $|EN'(i_q',m)|$, the least upper bound for any value of $m = m_1$ occurs at $(\hat{c}, m_1)$, where

$$(9) \qquad |EN'(\hat{c}, m_1)| \geq |EN'(c, m_1)| \text{ for all c.}$$

Upper bounds, which are not the least, are easily obtained as the envelopes of $|\rho_{dn}(m)|$, $|EN(m)|$ and $|DFM_k(m)|$. As $d_x(x)$ is an even function of x, all $DFM_k(m) \equiv 0$ for k odd. Hence upper bounds on odd k $DFM_k'(i_q',m)$ are not obtainable as envelope functions; they can be calculated by the $DFM_k'(c,m)$ method outlined above. This has not been done here.

In the following derivations, the subscript $_{\sim}(_{\Delta})$ will indicate that the dither used is sinusoidal (sawtooth).

2. SINUSOIDAL DITHER

A. Sinusoidal dither noise correlation coefficient $\rho_{dn\sim}$.

The variables involved are shown in Table 1,

Table 1

VARIABLES IN SINUSOIDAL DITHER

Range of θ	o_q	$n_q = o_q - i_q$	$d_\sim n_q$
$0 < \theta < \sin^{-1} \frac{1}{2m}$	0	$- d_\sim$	$- d_\sim^2$
$\sin^{-1}\left(\frac{1}{2m}\right) < \theta < \sin^{-1}\left(\frac{3}{2m}\right)$	q	$q - d_\sim$	$- d_\sim(d_\sim - q)$
$\sin^{-1}\left(\frac{3}{2m}\right) < \theta < \sin^{-1}\left(\frac{5}{2m}\right)$	2q	$2q - d_\sim$	$- d_\sim(d_\sim - 2q)$
⋮	⋮	⋮	⋮
$\sin^{-1}\left[\frac{(2B-1)}{2m}\right] < \theta < \sin^{-1}\left[\frac{(2B+1)}{2m}\right]$	Bq	$Bq - d_\sim$	$- d_\sim(d_\sim - Bq)$

where $i_q = d_\sim \equiv mq \sin \theta$, $\theta \equiv \omega_d t$, $B \equiv m$ rounded off to the nearest integer.

To obtain $\rho_{dn\sim}$, we have

$$(10) \quad (\sigma_n)^2_\sim = \frac{1}{2\pi} \int_0^{2\pi} n_q^2(\theta) d\theta$$

$$= q^2 \left\{\frac{m^2}{2} + B^2 - \frac{2}{\pi} \sum_{R=1}^{B} \left[(2R-1)\sin^{-1}\left[\frac{(2R-1)}{2m}\right] + \sqrt{4m^2 - (2R-1)^2}\right]\right\}$$

Also,

$$(11)\quad (\sigma_d)^2_{\sim} = \frac{q^2m^2}{2\pi}\int_0^{2\pi} \sin^2\theta d\theta = \frac{q^2m^2}{2},$$

and

$$(12)\quad \overline{(d_q n_q)}_{\sim} = \frac{1}{2\pi}\int_0^{2\pi} d_q(\theta)n_q(\theta)d\theta =$$

$$- q^2\left[\frac{m^2}{2} - \frac{1}{\pi}\sum_{R=1}^{B}\sqrt{4m^2 - (2R-1)^2}\right].$$

Therefore, we obtain:

$$(13)\quad \rho_{dn\sim}(m) = \frac{\left\{\frac{1}{\pi}\sum_{R=1}^{B}\sqrt{4m^2 - (2R-1)^2} - \frac{m^2}{2}\right\}}{\frac{m}{2}\sqrt{\frac{m^2}{2} + B^2 - \frac{2}{\pi}\sum_{R=1}^{B}\left[\sin^{-1}\left[\frac{(2R-1)}{2m}\right] + \sqrt{4m^2 - (2R-1)^2}\right]}}.$$

Using this equation, $\rho_{dn\sim}$ was plotted for m lying in the range $0.4 < m < 10$ (Fig. 3). Obviously, $\rho_{dn\sim} = -1$ for all $m \leq 1/2$.

By computer simulation it was verified that the envelopes of Figs. 3-9 are, indeed, upper bounds which are not always least. The examples in Appendix I illustrate this point.

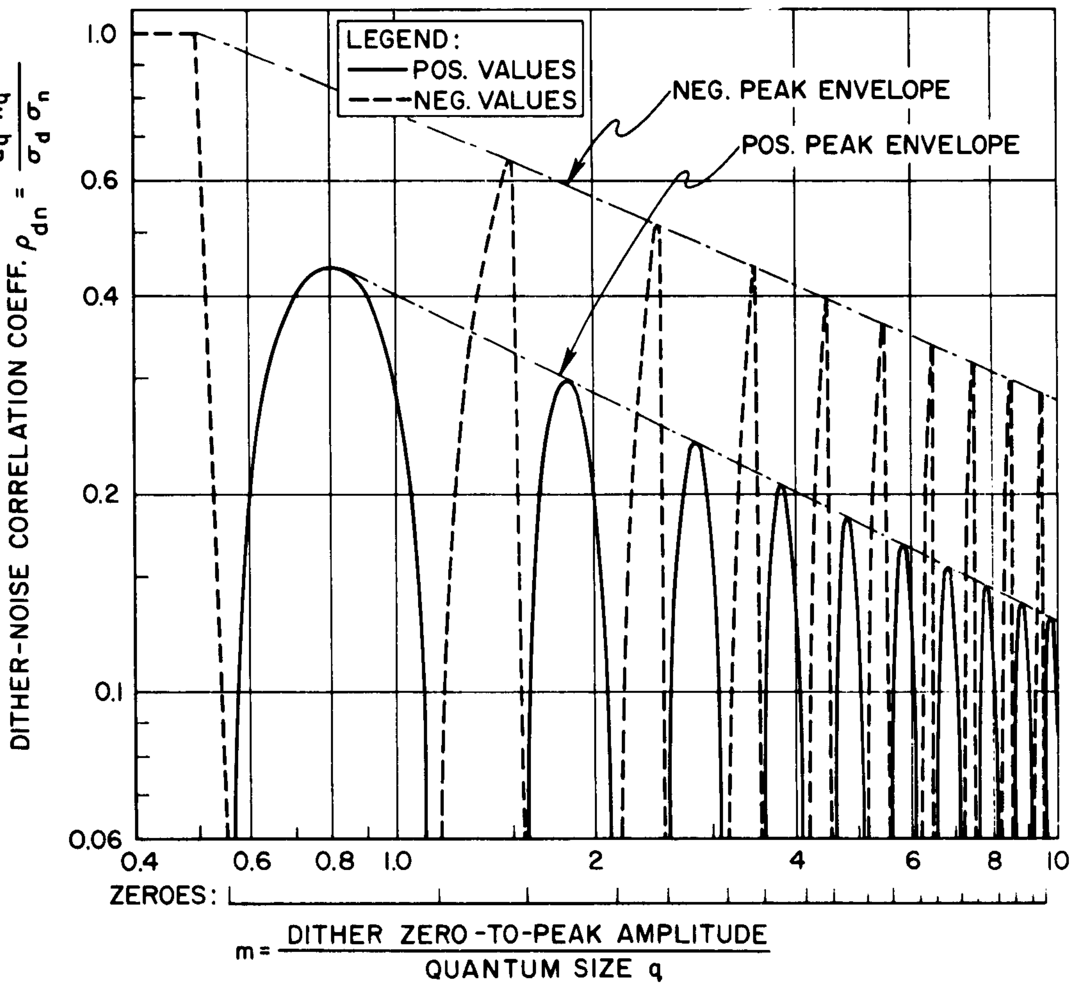

Fig. 3 -- Zero-shift correlation between d_q and n_q for sinusoidal dither

B. Sinusoidal dither figures of merit for moment fidelity $DFM_{k\sim}$.

Employing induction in the above manner we find that

$$(14)\quad DFM_{k\sim}(m) = -1 + \frac{B^k - \frac{2}{\pi}\sum_{R=1}^{B}[R^k - (R-1)^k]\sin^{-1}\left(\frac{2R-1}{2m}\right)}{\overline{[d_{\sim}(m) + n_{qu}]^k}}$$

for k even

= 0, for k odd.

Substituting k = 2, we obtain

$$(15)\quad DFM_{2\sim}(m) = -1 + \frac{B^2 - \frac{2}{\pi}\sum_{R=1}^{B}(2R-1)\sin^{-1}\left(\frac{2R-1}{2m}\right)}{\frac{m^2}{2} + \frac{1}{12}},$$

and for k = 4

$$(16)\quad DFM_{4\sim}(m) = -1 + \frac{B^4 - \frac{2}{\pi}\sum_{R=1}^{B}(4R^3 - 6R^2 + 4R-1)\sin^{-1}\left(\frac{2R-1}{2m}\right)}{\frac{3m^4}{8} + \frac{m^2}{4} + \frac{1}{80}}.$$

Figures 4 and 5 are computer generated plots of Eqs. (15) and (16) respectively.

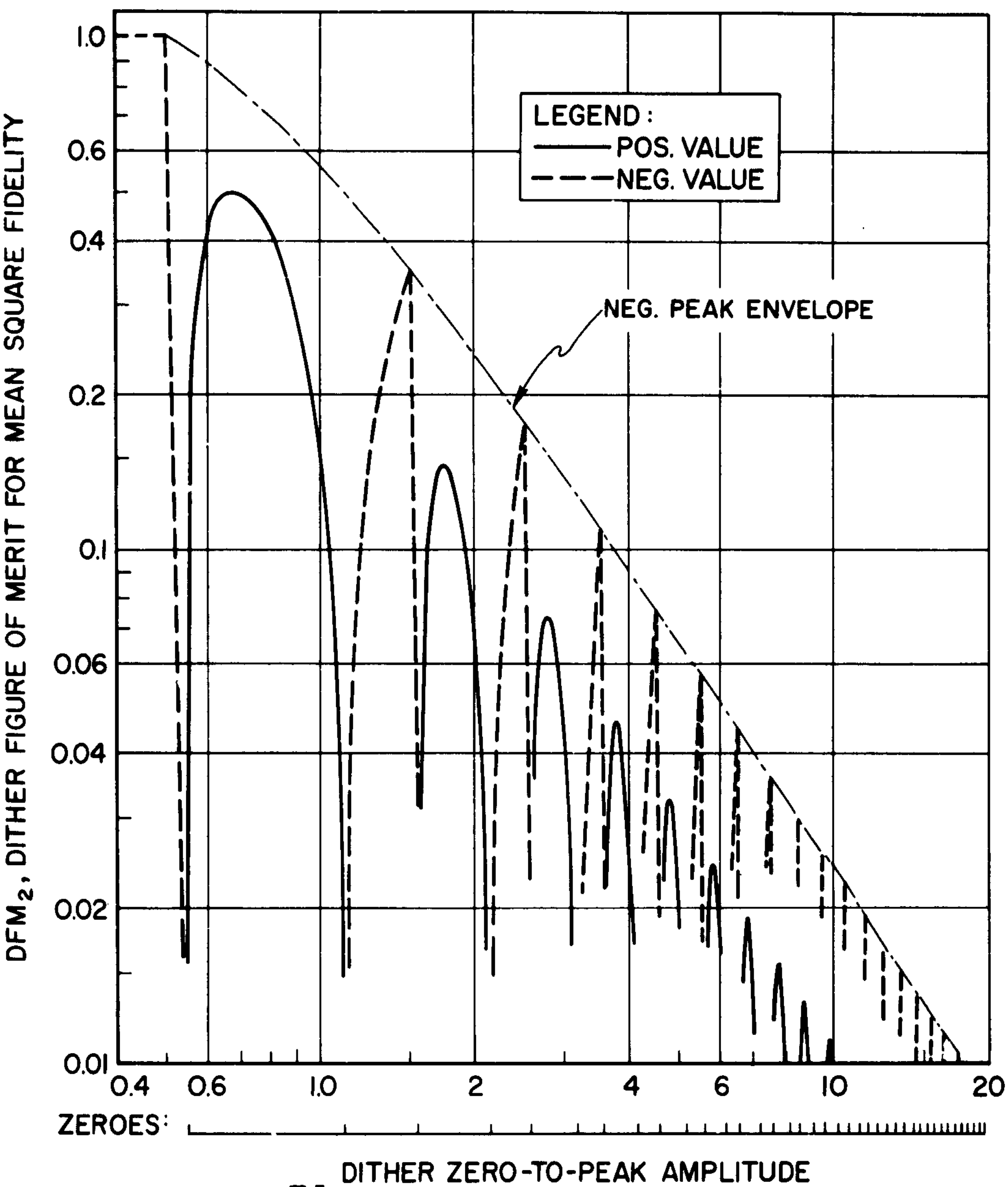

Fig. 4 -- DFM_2 for sinusoidal dither

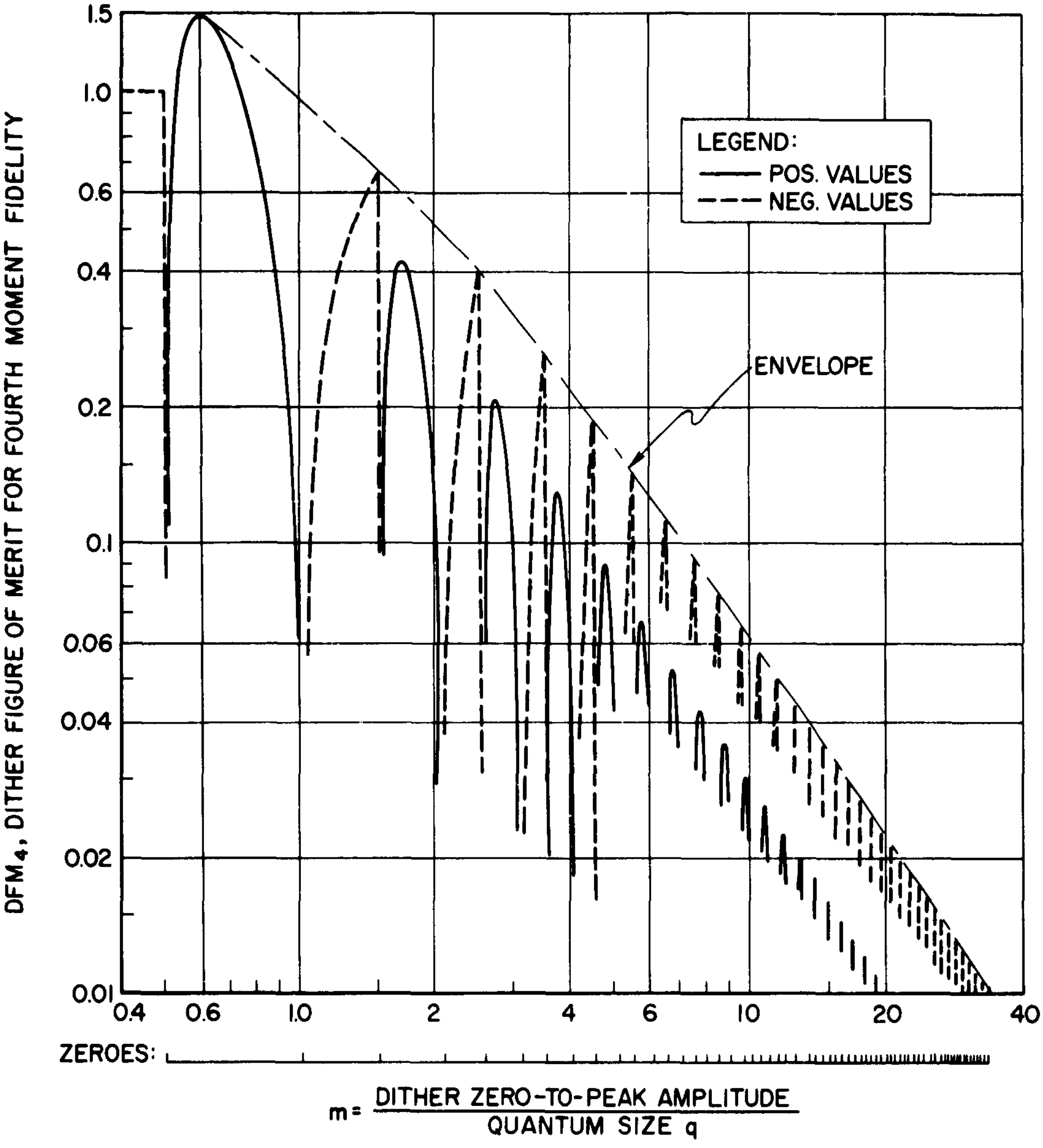

Fig. 5 -- DFM_4 for sinusoidal dither

C. The excess noise, $EN_{\sim}(m)$, for sinusoidal dither.

By employing Table 1, we obtain for the normalized mean square noise:

$$(17)\quad \frac{\overline{[n_q(m)]^2_{\sim}}}{q^2} = \left\{\frac{m^2}{2} - \frac{2}{\pi}\sum_{R=1}^{B}\left[\sqrt{4m^2 - (2R-1)^2} + (2R-1)\sin^{-1}\left(\frac{2R-1}{2m}\right)\right]\right\} .$$

In Fig. 6, a plot of Eq. (17) for the range $0.4 < m < 10$ shows the noise mean square to oscillate about its ideal value $n_{qu} = q^2/(12)$, the damping increasing with m. $EN_{\sim}(m)$ is obtained by combining Eqs. (4) and (17). Note that $\overline{(n_q)^2_{\sim}} = \overline{(d_{\sim})^2} = (mq)^2/2$ for $0 \leq m \leq 1/2$.

3. SAWTOOTH DITHER

A. For $i_q(t) = d_{\Delta}(t)$ we have the following results:

After some calculations (see Appendix II), we obtain for the noise mean square:

$$(18)\quad \overline{(n_q)^2_{\Delta}} = (1) = 2q^2\left\{\alpha^2 t_B + 2\alpha m t_B^2 + \frac{4m^2 t_B^3}{3} + 4m^2\left[\frac{1/8-\lambda^3}{3} - \frac{1/4-\lambda^2}{2} + \frac{1/2-\lambda}{4}\right]\right\}$$

for $1/2 \leq m \leq 3/2$.

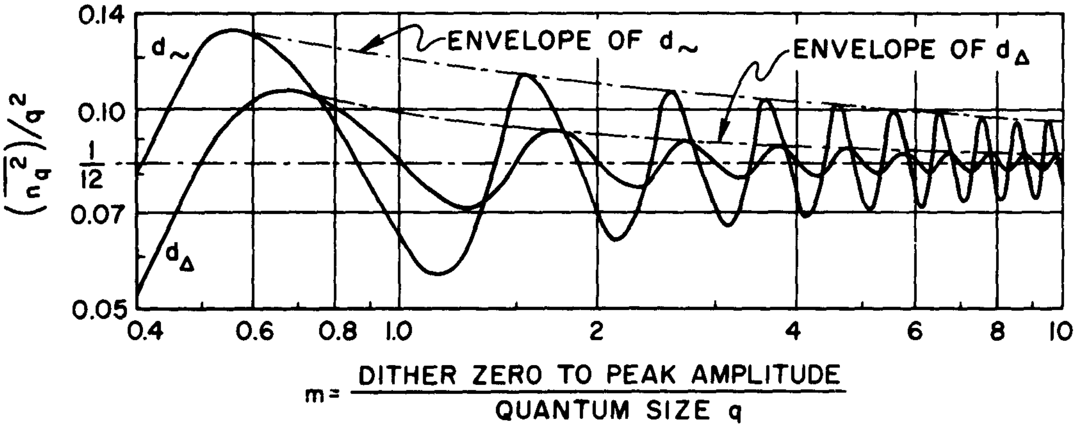

Fig. 6 -- Normalized noise mean square for $d_\sim$ and d_Δ

(19) $$\overline{(n_q)^2_\Delta} = 1 + 8q^2m^2 \sum_{R=1}^{B-1} \frac{\gamma^3-\delta^3}{3} - \epsilon(\gamma^2-\delta^2) + \epsilon^2(\gamma-\delta)$$

for $m \geq 3/2$

where $B \equiv m$ quantized, $\alpha \equiv B - m$, $t_B \equiv 1/2\left[1 - \frac{(2B-1)}{2m}\right]$,

$$\lambda \equiv \frac{(2m-1)}{4m}, \quad \gamma \equiv 1/2\left[1 + \frac{(1-2R)}{2m}\right], \quad \delta \equiv 1/2\left[1 + \frac{(1+2R)}{2m}\right]$$

and $\epsilon \equiv \frac{(m-R)}{2m}$.

Also, we obtain for the output k^{th} moment:

(20) $$\overline{(o_q)^k_\Delta} = q^k \sum_{R=1}^{B} \left[R^k - (R-1)^k\right]\left[1 - \frac{(2R-1)}{2m}\right]$$

for k even

= 0 for k odd.

In addition we obtain:

$$\overline{d^2_\Delta} = (mq)^2/3, \quad \sigma_{d\Delta} = \frac{mq}{\sqrt{3}} \quad \text{and} \quad \sigma_{n\Delta} = [\overline{n_q^{\,2}}]^{1/2} .$$

We obtain the desired expressions by writing:

$$\text{(21)} \qquad \rho_{dn\Delta} = \frac{(\overline{o_q^2} - \overline{d_\Delta^2} - \overline{n_q^2})}{2[\overline{d_\Delta^2}\ \overline{n_q^2}]^{1/2}}$$

and

$$\text{(22)} \qquad DFM_{k\Delta} = \frac{\overline{o_q^k} - \overline{(d_\Delta + n_{qu})^k}}{\overline{(d_\Delta + n_{qu})^k}} \quad \text{, for k even}$$

$$= 0, \text{ for k odd.}$$

Substituting $k = 2$ and $k = 4$ in the above expressions, we obtain:

$$\text{(23)} \qquad DFM_{2\Delta}(m) = \frac{\sum_{R=1}^{B} [2R-1]\left[1 - \frac{(2R-1)}{2m}\right] - \frac{m^2}{3} - \frac{1}{12}}{\frac{m^2}{3} + \frac{1}{12}}$$

and

$$\text{(24)} \qquad DFM_{4\Delta}(m) = \frac{\sum_{R=1}^{B} [4R^3-6R^2+4R-1]\left[1 - \frac{(2R-1)}{2m}\right] - \frac{m^4}{5} - \frac{m^2}{6} - \frac{1}{80}}{\frac{m^4}{5} + \frac{m^2}{6} + \frac{1}{80}} .$$

Using Eqs. (18) and (19), $\overline{(n_q)^2_\Delta}/q^2$ was plotted (Fig. 6) for the range $0.4 < m < 10$. Observe that

$\overline{(n_q)^2_\Delta} = \overline{d_\Delta{}^2} = (mq)^2/3$ for $0 \leq m \leq 1/2$.

Figures 7, 8, and 9 are plots of Eqs. (21), (23), and (24) respectively.

B. It was shown [1] that sawtooth dithers with $m = r/2$, where r is a positive integer, produce an equivalent quantizer gain which is unity. These dithers also have interesting properties from a statistical viewpoint, which are, as proven in Appendix III:

a. $EN_\Delta(m = r/2) = 0$.

b. For n odd we have: $\rho_{dn\Delta}(m = r/2) = -1/r$

and $DFM_{2\Delta}(m = r/2) = -2/(r^2+1)$.

c. For n even we have: $\rho_{dn\Delta}(m = r/2) = 1/(2r)$

and $DFM_{2\Delta}(m = r/2) = 1/(r^2+1)$.

4. CONCLUSIONS

A. Using the definition and measures of ideal quantization given in Sec. 1, we have determined numerically the departure of the actual quantization quality from the ideal for sinusoidal and sawtooth quantizer inputs. It was found

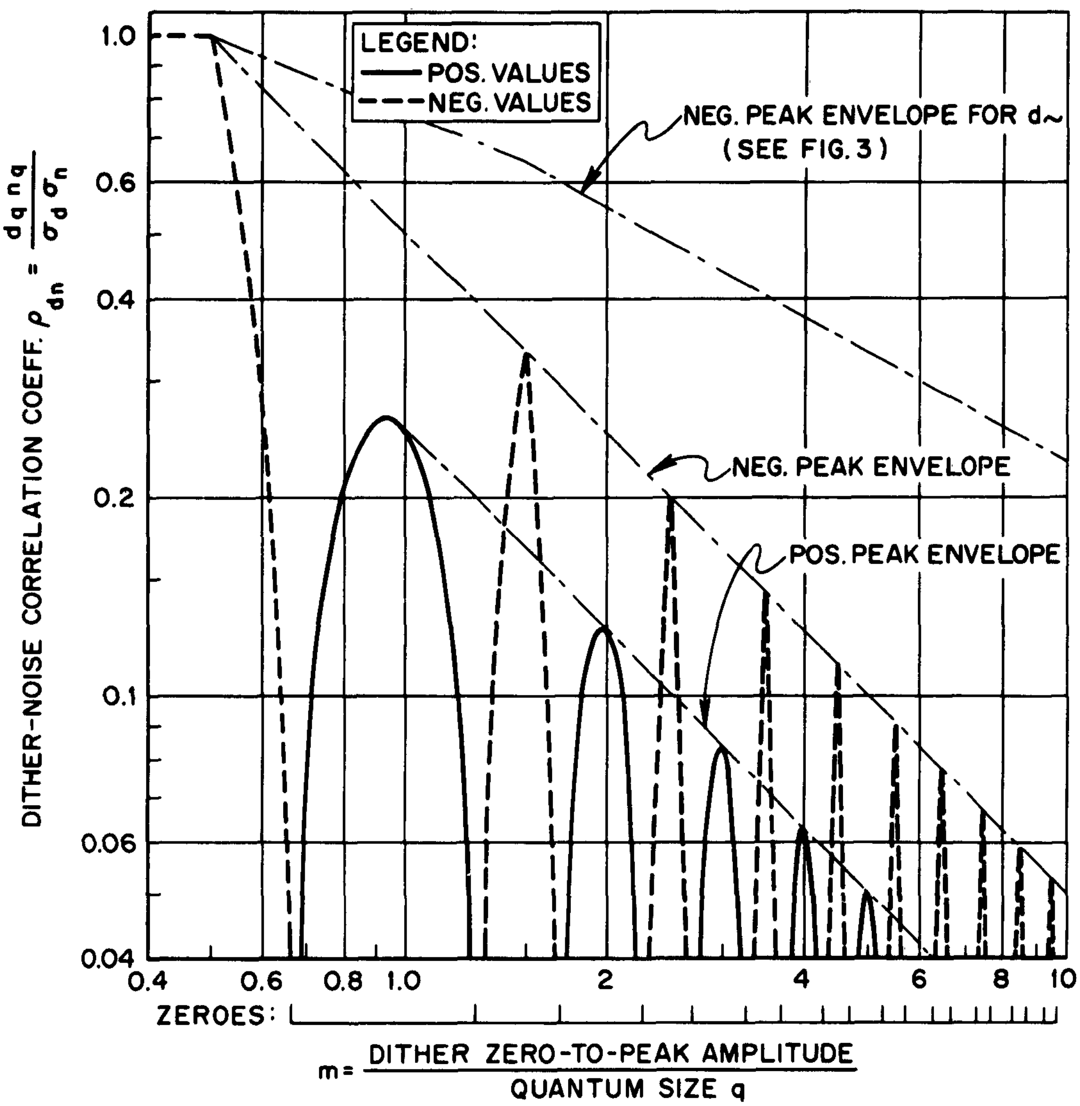

Fig. 7 -- Zero-shift correlation between d_q and n_q for sawtooth dither

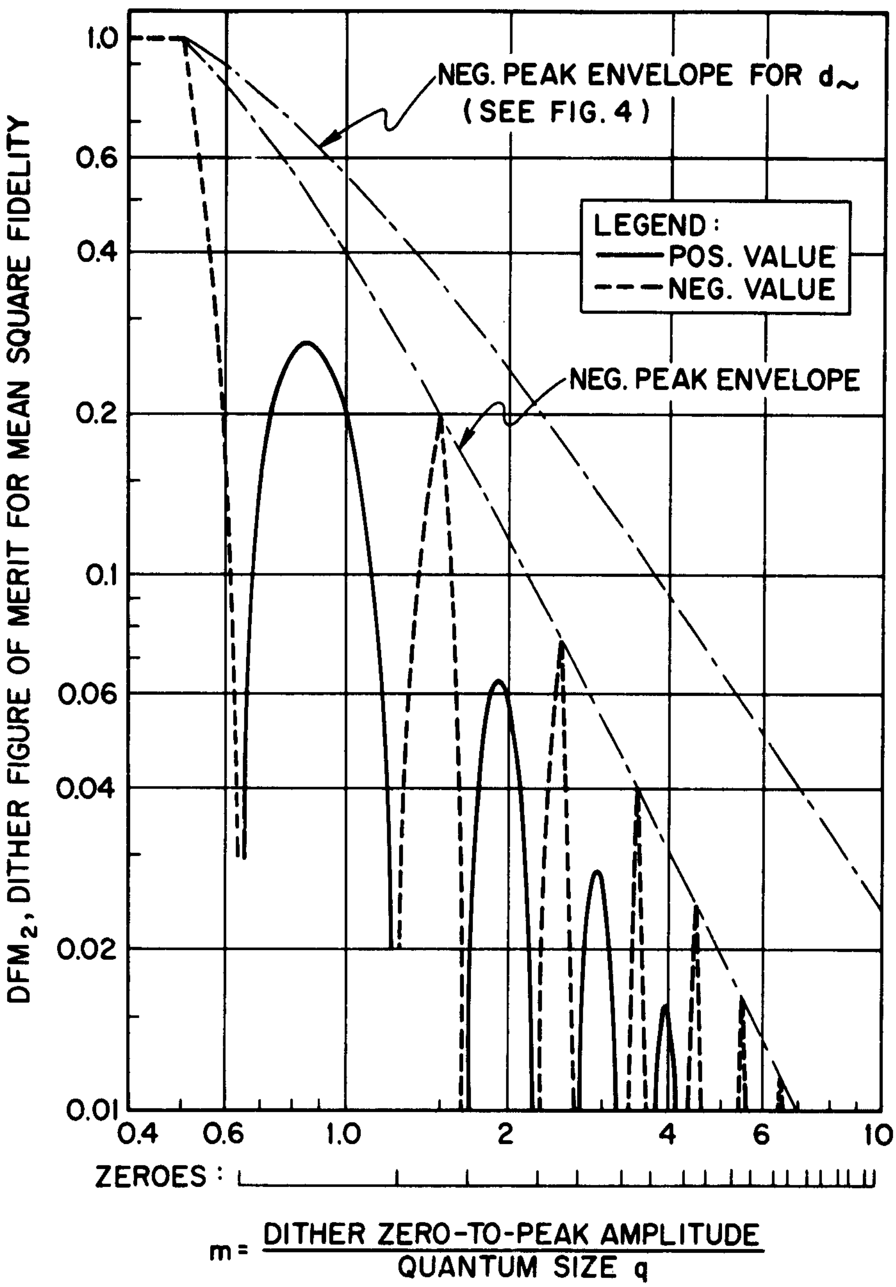

Fig. 8 -- DFM$_2$ for sawtooth dither

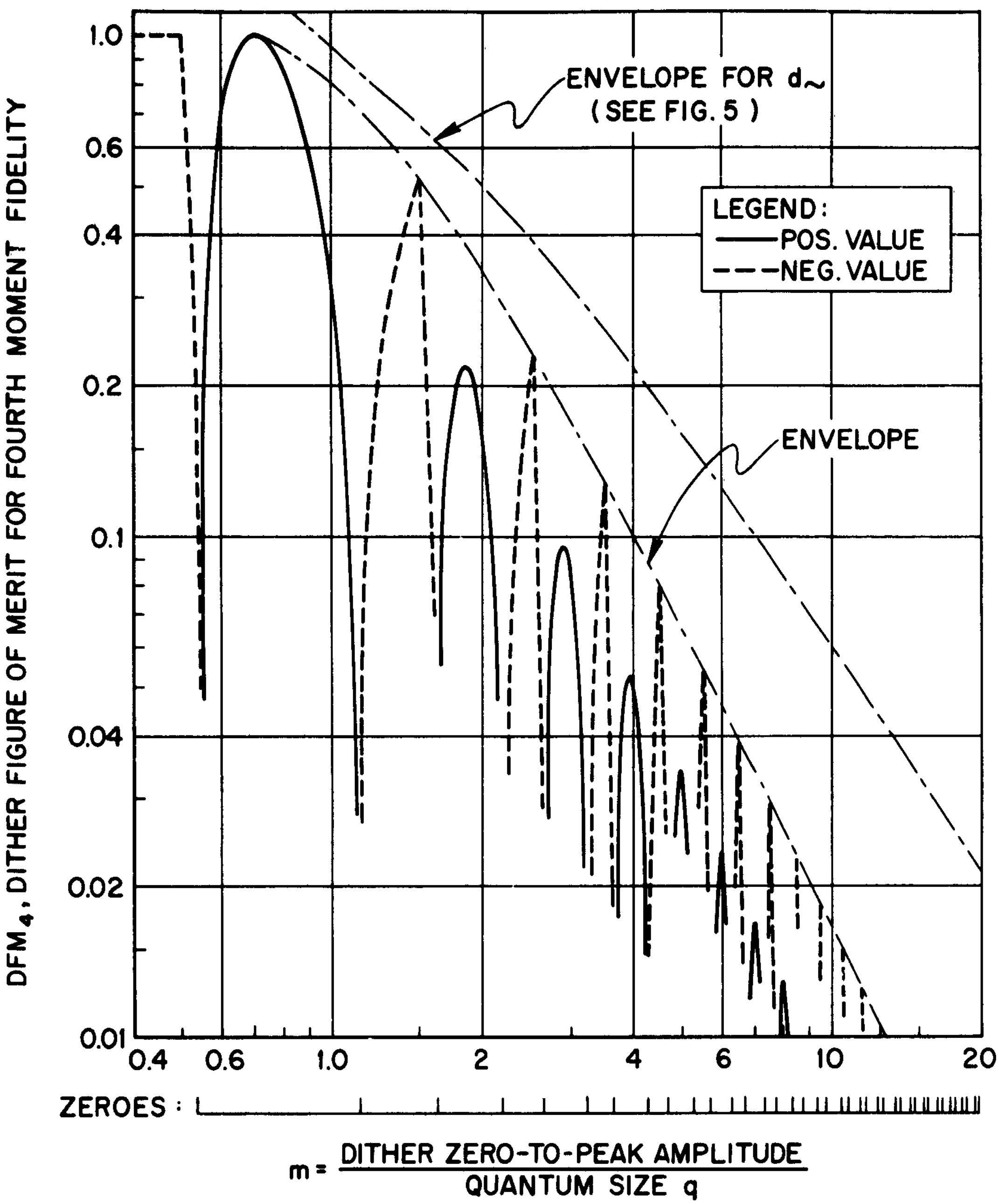

Fig. 9 -- DFM_4 for sawtooth dither

that any one of the three measures (Eqs. (2), (4), or (5)) can be made to zero (i.e., to be ideal) for certain values of m but that they cannot be made to zero simultaneously. For example, $EN_{\Delta}(1) = 0$ but $\rho_{dn\Delta}(1) = 1/4$, and $DFM_{2\Delta}(1) = 1/5$.

For both $d_{\sim}$ and d_{Δ}

$$(25) \qquad |\rho_{dn}(m)| = |DFM_2(m)| = |DFM_4(m)| = 1 \text{ for } m \leq 1/2.$$

However, for m > 1/2, the envelopes of the sawtooth functions lie below those for the corresponding sinusoidal functions. The results for sinusoidal and sawtooth inputs can be used to determine quantization quality directly in <u>undithered</u> systems, i.e., where $d_q(t) \equiv 0$ and $i_q'(t)$ is composed of sinusoids and/or sawtooths. These results, which appear in Secs. 2 and 3, constitute theorems for the quantization of these two fundamental wavetypes.

<u>B</u>. For a dithered system, i.e., where $i_q'(t) \neq 0$ and $d_q = d_{\sim}$ or d_{Δ} the envelopes of $|\rho_{dn}(m)|$, $|DFM_k(m)|$ and $|EN(m)|$ are of interest because they represent upper bounds for $|\rho_{in}'|$, $|DFM_k'|$, and $|EN'|$ for all i_q'.

These bounds, as given in Table 2, are conservative because they are not everywhere least even for the extreme situation where $i_q'(t) = c$. Nevertheless, they are useful for design purposes where it is desired to improve quantization

quality (and therefore some index of system performance such as the mean square error) to some specified level with $i_q'(t)$ unknown.

Table 2

FOR DESIGN PURPOSES

Measure	Dither Type	Range of m	Expression
ρ_{in}'	Sinusoid	$1/2 \leq m \leq 10$	$\lvert \rho_{in}' \rvert_{\sim} \leq 0.77\ m^{-0.43}$
ρ_{in}'	Sawtooth	$m \geq 1/2$	$\lvert \rho_{in}' \rvert_{\Delta} \leq \frac{1}{2m}$
DFM_2'	Sinusoid	$1/2 \leq m < 3/2$	$\lvert DFM_2' \rvert_{\sim} \leq 0.518 m^{-0.95}$
DFM_2'	Sinusoid	$3/2 \leq m < 20$	$\lvert DFM_2' \rvert_{\sim} \leq 0.63\ m^{-1.45}$
DFM_2'	Sawtooth	$m \geq 1/2$	$\lvert DFM_2' \rvert_{\Delta} \leq \frac{2}{4m^2+1}$
EN'	Sinusoid	$1/2 \leq m \leq 10$	$\lvert EN' \rvert_{\sim} \leq (0.127 m^{-0.123} - \frac{1}{12})$
EN'	Sawtooth	$1/2 \leq m < 5/2$	$\lvert EN' \rvert_{\Delta} \leq (0.103 m^{-0.1305} - \frac{1}{12})$
EN'	Sawtooth	$5/2 \leq m \leq 10$	$\lvert EN' \rvert_{\Delta} \leq (0.093 m^{-0.039} - \frac{1}{12})$

C. For dithers which abolish quantizer dead zone (i.e., for m > 1/2), sawtooth dither is superior to sinusoidal dither on all theoretical counts. This superiority has been demonstrated here on the basis of envelopes and in an earlier report [1] on the basis of degree-of-linearization. Hence, the only reasons for employing the sinusoid, its

higher m requiring greater quantizer capacity, are that it is sometimes more convenient to generate and/or inject.

D. As an example, let $m = 2$. The results appear in Table 3.

Table 3

COMPARISON OF DITHERS FOR $m = 2$

Operation	$\lvert \rho_{in}' \rvert$	$\lvert DFM_2' \rvert$	$\lvert DFM_4' \rvert$	$\lvert EN' \rvert$
Ideal	0	0	0	0
Sinusoidal dither	$\lvert \rho_{in}' \rvert_{\sim} \le 0.57$	$\lvert DFM_2' \rvert_{\sim} \le 0.23$	$\lvert DFM_4' \rvert_{\sim} < 0.52$	$\lvert EN' \rvert_{\sim} < 0.034$
Sawtooth dither	$\lvert \rho_{in}' \rvert_{\Delta} \le 0.25$	$\lvert DFM_2' \rvert_{\Delta} < 0.118$	$\lvert DFM_4' \rvert_{\Delta} < 0.35$	$\lvert EN' \rvert_{\Delta} < 0.0087$

With $m = 2$, the output mean square differs from its ideal value by less than 25% using the sinusoid but the difference is less than 12% using the sawtooth. Also the noise is more nearly ideal for the sawtooth as $\lvert \rho_{in}' \rvert_{\Delta\ max} < \lvert \rho_{in}' \rvert_{\sim\ max}$ and $\lvert EN' \rvert_{\Delta\ max} < \lvert EN' \rvert_{\sim\ max}$.

The output mean square is within 1% of its ideal value for all $m > 7$ using the sawtooth (Fig. 8), and for all $m > 18$ using the sinusoid (Fig. 4).

Appendix I

THE ENVELOPES OF $DFM_{2\sim}(m)$ AND $DFM_{2\Delta}(m)$

A. $DFM_{2\sim}(m)$. The purpose here is to show (for one value of m) how it was checked that the envelope of $|DFM_{2\sim}(m)|$ is an upper bound on $|DFM_2'(c,m)|_\sim$, where c is a constant function of time. Substituting k = 2 in Eq. (8), we obtain:

$$(26) \qquad |DFM_2'(c,m)|_\sim = \left| -1 + \frac{c^2}{\left(\frac{mq}{2}\right)^2 + c^2 + \frac{q^2}{12}} \right| .$$

The least upper bound is found for any value of m (say 5) by allowing c to vary freely. Equation (26) is plotted in Fig. 10 for $0.3q < c < 10q$ to include the maximum for m = 5 which occurs at $c \equiv \hat{c} = q/2$. Substituting, we obtain $|DFM_2'(c = q/2, m = 5)|_\sim = 0.0638$ as the least upper bound for m = 5, using sinusoidal dither. From Fig. 4 we read 0.065 for the envelope value at m = 5.

Some values of Eq. (26) for m = 0 were plotted together with those for m = 0.2 to illustrate the great improvement in $|DFM_2'|$ possible with a small value of m (Fig. 10). Let the sequence of maxima of $|DFM_2'(c,m)|_\sim \equiv S_{c,m} = S_{c,m}(1), S_{c,m}(2), \ldots S_{c,m}(n), \ldots$. Observe that $\lim_{n\to\infty} S_{c,0}(n) = 0$, as Fig. 10 implies, but that $S_{c,0}(n) > S_{c,0.2}(n) > S_{c,5}(n)$ for all n.

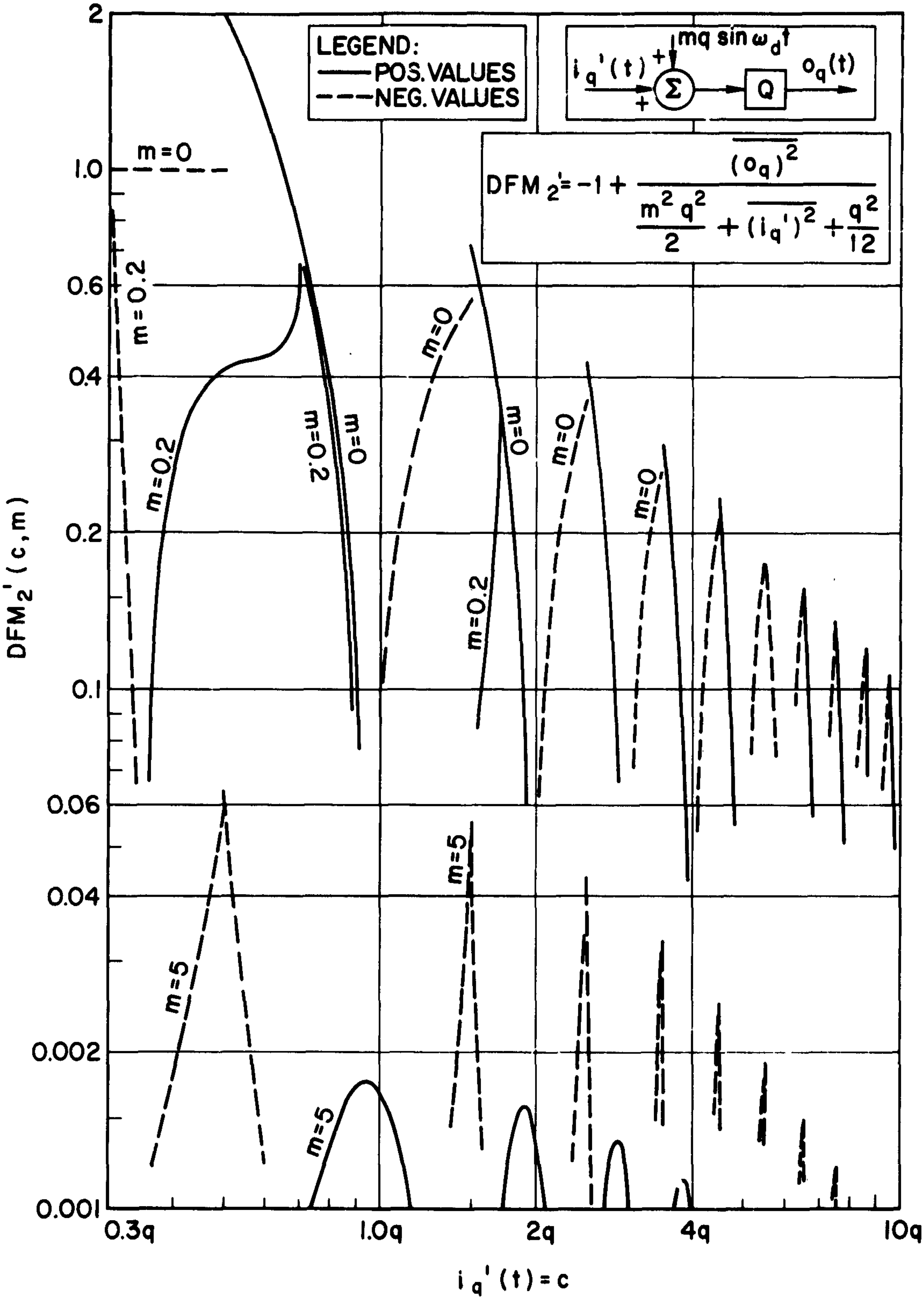

Fig. 10 -- Examples of DFM_2' (c, m)

(27) $$S_{c,0} = 2,\ 0.7,\ 0.44,\ \ldots\ ;$$

$$S_{c,0.2} = 1,\ 0.66,\ 0.37,\ \ldots\ ;$$

$$S_{c,5} = 0.064,\ 0.056,\ 0.044,\ \ldots\ .$$

B. $DFM_{2\Delta}(m)$. The purpose here is to illustrate that the envelope of $|DFM_{2\Delta}(m)|$, while an upper bound on $|DFM_2'(c,m)|_\Delta$, is not least for all values of m.

The envelope in question is $2/(4m^2 + 1)$. It is least at the maxima of $|DFM_{2\Delta}(m)|$ which occur at $m = 1/2,\ 3/2,\ 5/2,\ \ldots$, but it is not least elsewhere. Consider, for example, $m = 3/4$. By computer programming, one readily obtains $\hat{c} = q/2$ and the least upper bound is $3/5$, while the envelope gives a bound of $8/(13)$, which is 2.5% above the least.

Appendix II

DERIVATION OF $\overline{n_q^{\,2}}$ AND $\overline{o_q^{\,k}}$ FOR SAWTOOTH DITHER (i.e., $i_q = d_\Delta$).

A. $\overline{(n_q)_\Delta^2}$.

Let $t_i \equiv \frac{T_d}{2}\left(1 - \frac{2i-1}{2m}\right)$ and B = m quantized.

Then we obtain for

$$\overline{(n_q)_\Delta^2} = \frac{1}{T_d}\int_0^{T_d} [n_q(t)]^2 dt$$

the following:

$$(28)\quad \overline{(n_q)_\Delta^2}\Big|_{B=1} = \frac{2q^2}{T_d}\left\{\int_0^{t_1}\left[1 - m + \frac{2mt}{T_d}\right]^2 dt + \int_{t_1}^{T_d/2}\left[\frac{2m}{T_d}\left(t - \frac{T_d}{2}\right)\right]^2 dt\right\}$$

$$(29)\quad \overline{(n_q)_\Delta^2}\Big|_{B=2} = \frac{2q^2}{T_d}\left\{\int_0^{t_2}\left[2 - m + \frac{2mt}{T_d}\right]^2 dt + \int_{t_2}^{t_1}\left[\frac{2m}{T_d}\left(t - \frac{T_d(m-1)}{2m}\right)\right]^2 dt + \int_{t_1}^{T_d/2}\left[\frac{2m}{T_d}\left(t - \frac{T_d}{2}\right)\right]^2 dt\right\}$$

$$\vdots$$

$$(30)\quad \overline{(n_q)^2_\Delta}\Big|_{B=B} = \frac{2q^2}{T_d}\left\{\int_0^{t_B}\left[B-m+\frac{2mt}{T_d}\right]^2 dt + \left(\frac{2m}{T_d}\right)^2\left[\int_{t_B}^{t_{(B-1)}}\left[t-\frac{T_d(m-B+1)}{2m}\right]^2 dt\right.\right.$$

$$\left.\left.+\int_{t_{(B-1)}}^{t_{(B-2)}}\left[t-\frac{T_d(m-2+B)}{2m}\right]^2 dt + \ldots + \int_{t_1}^{T_d/2}\left[t-\frac{T_d}{2}\right]^2 dt\right]\right\}$$

where B is the positive integer and $B - 1/2 \leq m \leq B + 1/2$.

Writing Eq. (30) in closed form, we obtain Eqs. (18) and (19).

B. $\overline{(o_q)^k_\Delta}$.

For k = 2, we obtain

$$(31)\qquad (\overline{o_q^2}/q^2)_\Delta\Big|_{B=1} = \left(1-\frac{1}{2m}\right)$$

$$(32)\qquad (\overline{o_q^2}/q^2)_\Delta\Big|_{B=2} = 3\left(1-\frac{3}{2m}\right) + \left(1-\frac{1}{2m}\right)$$

$$\vdots \qquad\qquad \vdots$$

$$(33)\qquad (\overline{o_q^2}/q^2)_\Delta\Big|_{B=B} = \sum_{R=1}^{B}\left[R^2-(R-1)^2\right]\left[1-\frac{(2R-1)}{2m}\right] .$$

Similarly, for $k = 4$ we obtain

$$(34)\qquad (\overline{o_q{}^4}/q^4)_\Delta = \sum_{R=1}^{B} [R^4 - (R-1)^4]\left[1 - \frac{(2R-1)}{2m}\right]$$

and in general

$$(35)\qquad (\overline{o_q{}^k}/q^k)_\Delta = \sum_{R=1}^{B} [R^k - (R-1)^k]\left[1 - \frac{(2R-1)}{2m}\right] .$$

Appendix III

PROPERTIES OF SAWTOOTH DITHERS WITH $m = r/2$

We are given

$$d_\Delta(t) = \frac{qr}{2}\left(1 - \frac{2t}{T_d}\right) \qquad 0 < t < T_d$$

where r is a positive integer. Then obviously the quantization noise $n_q(t)$ is time periodic with period $T_n = T_d/r$. Also, $\sigma_d = qr/(2\sqrt{3})$ and $\sigma_n = \sigma_{nu} = q/(2\sqrt{3})$.

A. Odd r. We have

$$(36) \qquad n_q(t) = -\frac{q}{2}\left(1 - \frac{2rt}{T_d}\right) \qquad 0 < t < T_d/r$$

and

$$(37) \qquad \overline{d_\Delta n_q} = \frac{1}{T_d}\int_0^{T_d} d_\Delta(t)n_q(t)dt$$

$$= -\frac{r}{T_d}\int_0^{T_d/r}\left[\frac{q}{2}\left(1 - \frac{2rt}{T_d}\right)\right]^2 dt = -\frac{q^2}{12}\,.$$

Therefore,

$$(38) \quad \rho_{dn\Delta}(m = r/2) = \frac{1}{T_d\sigma_d\sigma_n}\int_0^{T_d} d_\Delta(t)n_q(t)dt = -1/r$$

for r odd.

Expanding Eq. (22) we obtain:

$$(39)\quad DFM_{2\Delta}(m=r/2) = \frac{\overline{d_\Delta^{\,2}} + \overline{2d_\Delta\, n_q} + \overline{n_q^{\,2}} - \overline{d_\Delta^{\,2}} - \overline{2d_\Delta n_{qu}} - \overline{n_{qu}^{\,2}}}{\overline{d_\Delta^{\,2}} + \overline{2d_\Delta n_{qu}} + \overline{n_{qu}^{\,2}}} .$$

By definition, $\overline{d_\Delta n_{qu}} \equiv 0$. Also, $\overline{n_q^{\,2}} = \overline{n_{qu}^{\,2}} = q^2/(12)$ and $\overline{d_\Delta^{\,2}} = (qr)^2/(12)$. Therefore,

$$(40)\quad DFM_{2\Delta}(m=r/2) = \frac{\overline{2\; d_\Delta n_q}}{\overline{d_q^{\,2}} + \overline{n_{qu}^{\,2}}} = \frac{24\overline{d_\Delta n_q}}{q^2(n^2+1)}$$

Therefore, we have

$$(41)\quad D^{-}M_{2\Delta}(m=r/2) = - \frac{2}{(n^2+1)}$$

for r odd.

<u>B</u>. Even r. We have

$$(42)\quad n_q(t) = \frac{rqt}{T_d} \qquad \text{for} \qquad -\frac{T_d}{2r} < t < \frac{T_d}{2r}$$

and

$$(43)\quad \overline{d_\Delta n_q} = \frac{1}{T}\left\{2\int_0^{T_d/(2r)}\left[\frac{qrt}{T_d}\right]\left[\frac{qr}{2}\right]\left[1-\frac{2t}{T_d}\right]dt-(r-1)\int_0^{T_d/r}\left[\frac{q}{2}\left(1-\frac{2rt}{T_d}\right)\right]^2 dt\right\}$$

$$= \frac{q^2}{24}\ .$$

Also, substituting in Eq. (40), we obtain

$$(44)\quad DFM_{2\Delta}(m=r/2) = \frac{1}{(n^2+1)}$$

for r even.

Substituting in Eq. (2) we obtain

$$(45)\quad \rho_{dn\Delta}(m=r/2) = \frac{1}{2r}$$

for r even.

REFERENCES

1. Furman, G. G., Removing the Noise from the Quantization Process by Dithering: Linearization, The RAND Corporation, RM-3271, January 1963.

2. Bennet, W. R., "Spectra of Quantized Signals," Bell System Technical Journal, Vol. 27, 1948, pp. 446-472.

3. Widrow, B., "Statistical Analysis of Amplitude-Quantized Sampled-Data Systems," AIEE Transactions, Vol. 79, Part II, Applications and Industry, 1961, pp. 555-568.